

Paper 102-29

Good data makes good marketing: Using data management to enhance the effectiveness of database marketing campaigns

John Walker, Bell Mobility, Mississauga, ON

Gavin Day, DataFlux Corporation, Cary, NC

ABSTRACT

The goal of database marketing should be to build relationships. When you create stronger, tighter relationships with your customers, they are likely to buy more often, purchase at a higher price, and try new products when they become available. Unfortunately, poor data quality hinders many database marketing efforts. If your customer information contains incorrect, duplicate, or incomplete data, you cannot effectively reach your entire universe of contacts—and can waste time and resources contacting the wrong address or phone number.

INTRODUCTION

This paper provides a real-world example of the technology and techniques required to verify, merge and cleanse customer data prior to starting database marketing efforts. By performing extensive data quality routines on new and existing customer contact information, companies can provide more reliable data to its telemarketing and direct marketing groups—and allow them to contact more people than in a traditional database marketing campaign. As a result, database marketing efforts can contact more prospective customers. And help you elevate the effectiveness of marketing campaigns.

DATABASE MARKETING DEFINED

Database marketing is the process of using data to deliver marketing messages to specific target markets. As you refine the message, you can more effectively change the behavior of each market. And the more specific you can be about the market you intend to reach, the less it can cost to successfully reach each member of that market.

A fundamental element of database marketing is to learn as much as you can about prospects and customers, and reflect that knowledge in marketing activities. When you elevate the level of marketing communications with target markets, you are well on your way to building lasting and profitable customer relationships, turning prospects into customers and customers into loyalists. And, when you have close relationships with customers, they are more likely to buy more often, purchase at a higher price, and try new products.

The ultimate in database marketing is targeting a specific message to a target market of one. Imagine receiving a telephone call that goes like this:

“Good morning, <your name>. I’m calling from ABC-XYZ Distributors, and we’re glad to have you as a valued client. While reviewing your records recently, we noticed that your company has purchased \$150,000 in supplies from us over the past 6 months . Based on your total volume of business with us, and that you’ve been a customer for five years, we can show you how you can reduce shipping charges and decrease per unit costs by consolidating your company’s activities into one account. In the process, you won’t lose any flexibility in buying supplies, and you will gain access to a package of additional services, such as online bill pay, at no cost to you. I understand that you’re usually quite busy during the workweek, but I would be happy to schedule a time to talk with you to discuss the details. Is this something that interests you?”

With a sophisticated marketing database, the information in the systems is used to develop a more effective message. And, in this example, you can see how database marketing can go beyond just managing name and address information. With better data management procedures, database marketing goes from being a communication method to a way of developing a 360-degree view of the customer, their preferences, their activities and, ultimately, their value to your organization.

But the foundation for this form of marketing is high quality data. To understand your customer base, you have to have a set of processes in place to build better data sources from the existing information you have. This methodology is commonly known as data management.

THE ROLE OF DATA MANAGEMENT IN MARKETING CAMPAIGNS

Marketers typically spend a large portion of time on the more creative parts of the marketing spectrum: designing an appealing look-and-feel, developing messaging that resonates with the target audience, testing messages to focus groups, and so forth. Marketers even spend days—and sometimes weeks—trying to decide exactly what type of individual meets the requirements of their target audience.

What is often left out of this process is inspecting, correcting, integrating and enhancing customer or prospect contact data. If your target audience doesn't actually receive the message, all the creative work spent to develop the message is useless. Or if they receive multiple copies of the same piece, customers and prospects can view your outbound communications as amateurish or, worse yet, annoying.

To make the most of your carefully-crafted marketing messages, you need to have consistent, accurate and reliable sources of contact information as the foundation for marketing programs. With more powerful, usable databases, you know that your recipients will receive the intended message.

THE EFFECT OF DATA WITHIN THE MARKETING FUNCTION

Transforming ineffective marketing databases (i.e., those that contain inconsistent or incorrect data on your customers and prospects) into more effective ones can greatly enhance your marketing efforts and save your organization considerable time and money. This paper outlines the common pitfalls of marketing databases and provides a real-world example of the effectiveness of data management within the marketing function.

Effective marketing databases—databases that contain usable, reliable data on customers and prospects—are often the most important elements of effective marketing campaigns. When you use good data to deliver marketing messages customized to each prospect and/or customer, the results can be phenomenal.

For any campaign, good information on prospects and customers means that you can more accurately group or segment recipients into different groups. This can boost response rates, improve customer retention, enhance customer satisfaction, increase referrals and cross-sales, drive up sales revenue, save money, and help build brand identity. When you use bad or incorrect data to deliver messages, the results can be disastrous. Inconsistent data can prevent messages from being delivered as intended, confuse recipients, erode customer loyalty, waste money, and cause you to miss valuable opportunities.

Many organizations have a wealth of marketing data about their customers and prospects. This information comes from a variety of sources over time. Sometimes, data comes from internal groups such as sales, marketing, finance/accounting and customer support. You might also receive data from data vendors, partner organizations, government agencies, and the Internet. Each data set can be unique, with different formats, protocols and structures—making it difficult to merge and assimilate different sources.

With data coming from a variety of sources, the potential for failure due to bad data is a real, ongoing problem for database marketers. The following table compares typical pitfalls of using weak data with some benefits of using strong data. As a whole, these pitfalls and benefits can negatively or positively affect costs, opportunities, response rates, customer satisfaction and retention, sales revenue, and brand identity.

Pitfalls of Bad Data	Benefits of Strong Databases
Incorrect and incomplete addresses...	Correct and complete addresses...
...prevent marketing messages from reaching everyone on your list. Example: The message reaches someone who currently lives at the customer's previous address, or the message is never delivered to anyone.	...help marketing messages reach everyone on your list.
...reduce the value of campaign tracking statistics. Example: Your campaigns normally have a response rate of 1%, but the latest campaign produced only .2%. Was the problem with the message itself or the appropriate people not receiving the message?	...increase the value of campaign tracking statistics. Example: If campaigns normally have a response rate of 1%, but the latest campaign produced only .2%. You know the problem was with the message itself, not the delivery of that message.
...waste production/delivery costs. Example: Each message that doesn't reach the intended recipient is a waste of money.	...make efficient use of production/delivery costs. Example: You mail 100,000 postcards. Of these, 99,997 people in the target market receive a postcard.
...miss opportunities to increase business by attracting new customers, preventing customer attrition, cross-selling, and up-selling.	...maximize opportunities to increase business through attracting new customers, preventing customer attrition, cross-selling, and up-selling.

Pitfalls of Bad Data	Benefits of Strong Databases
Incorrect prospect/customer name data...	Correct prospect/customer name data...
...mismatch messages and target markets. Example: A message about commercial tree fertilizers is delivered to John Orchard, an accountant.	...match messages and target markets. Example: A message about commercial tax software is delivered to John Orchard, an accountant.
...can annoy recipients with misspelled names. Example: John E. Mitchell receives the message, addressed to Jon Mitch.	...mean correctly spelled names.
...can cause incorrect assumptions about gender. Example: Mr. Pat Smith receives the message about lipstick, addressed to Ms. Pat Smith.	...usually lead to accurate assumptions about gender. Example: Mr. Pat Smith receives the message about shaving cream, while his wife, Terry, receives the message about lipstick.
Duplicate records...	Unique records...
...lead to duplicate messages. Example: One prospect receives 12 coupons for a free investment consultationhelp prevent duplicate messages. Example: Each prospect receives exactly one coupon for a free investment consultation.
...miss opportunities to accurately augment records. Example: You buy demographic data about Jim Johnson, but can't match the data to your records for James Johnson and Jimmy Johnson. They are all the same person.	...take advantage of opportunities to accurately augment records with additional data. Example: You buy demographic data about Jim Johnson, match the data to your record for Jim, and now have information about his age and income.
...prevent you from knowing the true value of a customer. Example: Margaret Smith spends \$1,000 on products each year. Meg Smith buys \$3,000. M. Smith buys \$5,000. Who is the most valuable customer? If they share common data, such as an address and/or telephone number, they are probably the same woman.	...help you know the true value of a customer. Example: Margaret Clark spends \$9,000 on products each year, even though she pays with three different checking accounts in three different names.

THE FOUR BUILDING BLOCKS OF DATA MANAGEMENT

To increase the value of your marketing database, one approach to data management within the database-marketing arena is divided into four main areas of focus, or "building blocks." These four building blocks are: Profiling, Quality, Integration, and Augmentation, as shown in Figure 1.

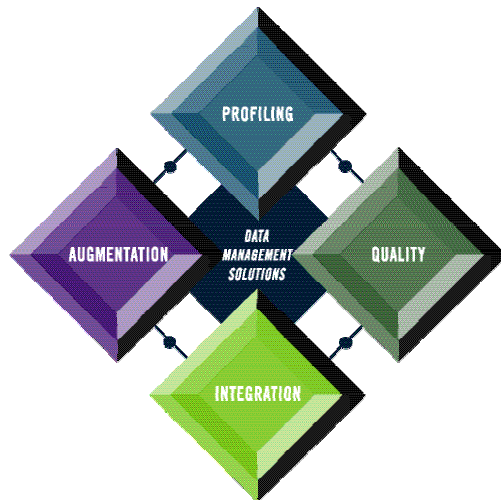


Figure 1: The Four Building Blocks of Data Management

DATA PROFILING

The first building block, data profiling, encompasses discovery and audit activities that help you assess the composition, organization, and quality of databases. Data profiling activities are an invaluable start for a data management cycle. These activities help you:

- Better understand existing databases
- See how well current databases might support potential marketing activities
- Determine what steps you should take to address data problems

For example, a profiling analysis might indicate that 80% of prospects and customers are males, 40% of the records are missing street addresses, and the databases have no income data. Suppose you are marketing a line of high-end women's shoes and accessories. After data profiling, you would know that this data set is not one that you could fully leverage to drive sales for these products.

You can also perform data profiling activities again at the end of each cycle to measure improvement and, through segmentation and numeric analysis, to gain new insights about prospects and customers. Data profiling tactics can also be used to validate information as it flows into your system. For instance, if your Web site collects customer information, you can periodically check the new information to see if the cumulative quality of data degrades as more information enters the system. This could cause you to restart a comprehensive data management process.

Data profiling tasks typically target three different attributes of data, each looking at the content from a different perspective:

- Structure analysis—Is the table structure accurate? Does the metadata support the data? Does the data conform to expected patterns?
- Data discovery—Is the data complete and usable? Is the data accurate and unambiguous?
- Relationship discovery—Does the data adhere to key relationships? Is there redundant data across tables?

To illustrate the use of data profiling within database marketing, suppose you buy a list of names from a list broker to augment your prospect database. Typically, you would send mailings to both lists and hope for the best response. If prospects are on both lists, however, it is a waste of money for them to receive mail twice. With a redundant data analysis, you can quickly compare two sources and receive the following report, provided by DataFlux's dfPower Studio[®] technology.

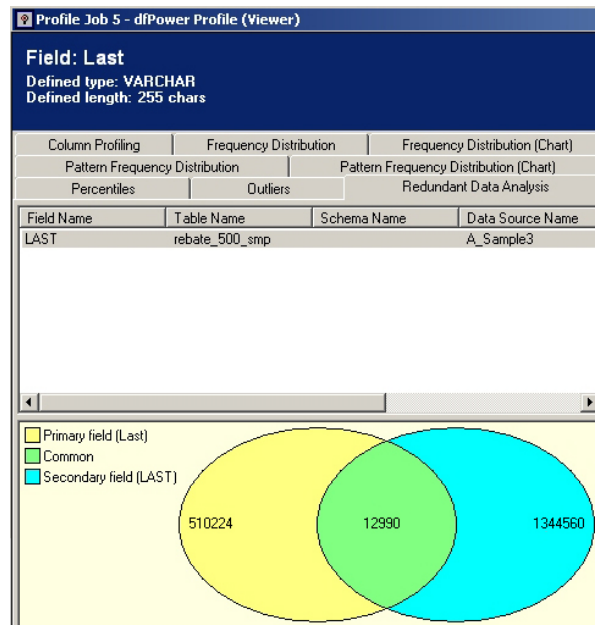


Figure 2. Redundant data analysis.

Through this analysis, you see that there are 12,990 names that overlap the two databases (in dfPower Studio, you can also click on the green portion to view the records that overlap). With this analysis, you know that before you send out a marketing message to the two databases, you could save time and money by building rules to remove duplicates before the communication occurs. Data profiling gives you the information to make these choices; the next phase gives you the tools to cleanse data.

DATA QUALITY

The second building block, data quality, encompasses activities for correcting, updating and validating existing data. Often referred to as data hygiene or cleansing, these activities allow you to take the information gleaned from the data profiling stage and build rules to fix data.

For example, if profiling shows that 80% of records are complete, but only 10% are less than two years old, your first priority may be to update the records. Or, the analysis shows that only 25% of mailing addresses are valid, but 75% of the phone numbers are verifiable. You would know that to do a direct mail campaign, you would have to work diligently to correct mailing addresses in order to reach the intended recipients. If you want to start reaching the contacts in this database more quickly, you may choose to update telephone numbers first and do a telemarketing campaign initially while correcting the mailing addresses.

Data quality routines can accomplish a number of tasks. One method, data standardization, helps you address and correct multiple permutations of data. The company ACME Manufacturing Corporation may be represented in the same data source as Acme Mftg Corp, ACME, and ACME Manufacturing. With DataFlux technology, you can initiate a scheme to change data to a standardized format. Once completed, you can get an accurate picture about the size of your total business relationship with ACME Manufacturing Corporation, because all permutations have now been standardized on one naming convention.

Similarly, pattern standardization helps you create valid patterns of data across tables and columns. Some pieces of data, such as phone numbers, product data or Social Security Numbers, may have a common pattern such as (999) 999-9999. Pattern standardization can take information in non-standard formats, such as 9999999999 or 999.999.999 and turn that into an accepted telephone number.

Another quality feature is householding—grouping individuals or organizations within an established group based on a common feature (usually the same address). Householding lets you leverage data about customers' relationships to each other to gain insight into their spending patterns, job mobility, family moves, additions, and more. This insight can help leverage up-sell or cross-sell opportunities based on family or organizational interests and needs.

In addition to improving the accuracy and reliability of databases, data quality activities often serve as the foundation for more sophisticated integration and augmentation activities. Following the quality phase, you have corrected a number of errors within your data. Now, the goal is to integrate information across your existing data sources.

DATA INTEGRATION

The third building block, data integration, includes activities for merging, linking/joining, and determining relationships among existing data. You can integrate data at multiple levels, or you can merge new data—such as purchased lists—into your prospect database and remove duplicate records.

Integration activities help you avoid waste and embarrassment by reducing or eliminating duplicate messages. Through data consolidation, these activities can also help you determine the most accurate and complete data for prospects and customers. For example, if a customer is recorded in different sources as Jonathan Smith with no street address, Jon Smith with an invalid telephone number, and J. Smith with no purchasing data, integration activities can build a single, more accurate view of Mr. Smith.

Data integration activities are also important when attempting to link or join databases. Imagine trying to join a purchased prospect database and your current customer database without a sophisticated set of matching algorithms to make sense of the potential overlap of data between the two. Without sophisticated matching, you might try joining records based on exact name and address matches, but this is likely to be much less fruitful than using match codes.

Consider the following records, each of which is in a separate data source:

Customer Name	Street Address	ZIP
Taylor C. Jones	212 Elm St.	95503-7936
Taylor Jones	212 E Elm St	95503
TC Jones	212 Elm Street	95503-7936

It looks like these three records are for the same customer. To determine this quickly and easily, a technology with sophisticated matching definitions, techniques, and algorithms can identify clusters of duplicate and near-duplicate records. To accomplish this, the technology can *parse* the data internally—for example, Customer Name would be parsed into first, middle, and last names—and then use fuzzy logic to assign a *match code* to each record, as shown in the following table:

Customer Name	Street Address	ZIP	Match Code
Taylor C. Jones	212 Elm St.	95503-7936	NNGE\$\$RTH&JK\$\$
Taylor Jones	212 E Elm St	95503	NNGE\$\$RTH&JK\$\$
TC Jones	212 Elm Street	95503-7936	NNGE\$\$RTH&JK\$\$

Even though the data in each record is a little different, the match codes are all the same. So, the technology considers them to be duplicates or near-duplicates.

The examples above show relatively conservative matches. While the customer names are different, the street addresses and ZIP codes are essentially the same, so it's reasonable for the technology to assume that the three records are for the same customer. Depending on your own data, you can adjust the technology to be more or less conservative by adjusting the *match sensitivity*. A high sensitivity requires closer matches and yields fewer of them. A low setting finds looser matches, and more of them.

DATA AUGMENTATION

The fourth building block, data augmentation, encompasses activities for supplementing marketing databases with external data and services. For example, you might use a database of typical male and female names in conjunction with sophisticated data parsing technology to make accurate guesses about the gender of each prospect and customer. Or, you might use a demographics database to identify revenue, number of employees and Standard Industrial Classification (SIC) codes for organizations. All of these pieces of data shed more light on the purchasing characteristics of your prospects and clients. You might also choose to use suppression information from a third-party vendor or from a government agency so you can be sure to meet regulatory requirements of "do-not-call" laws.

After conducting augmentation routines, you have enhanced your existing data, allowing you to better understand and predict the behavior of customers and prospects, and better personalize marketing messages to them. While data augmentation can take many forms, three common uses are:

- Suppression—for example, not sending offers for home-improvement services to the deceased, college students, apartment dwellers, people living in retirement homes, and anyone on the DMA no-mail list
- Address hygiene—updating and correcting your data via USPS -sanctioned address correction and enhancement processes.
- Business and consumer data appends—augmenting prospect and customer records with data about spending habits, lifestyle indicators, occupational codes, and family size

Let's look at using suppression data with the following consumer records:

First Name	Last Name	Street Address	ZIP
Bill	Jones	1324 Milton Road	27712
Mary	Wise	115 Dublin Woods Drive	27513
Jim	Phelps	1406 Oakcreek Drive	14450

Suppose you are marketing a program for new homebuyers. Using a third-party suppression database, we'll append these records with an additional field: Dwelling Type. This will help you decide if there are records that do not fall into the target audience.

First Name	Last Name	Street Address	ZIP	Dwelling Type	Do Not Mail
Bill	Jones	1324 Milton Road	27712	Single-family house	Y
Mary	Wise	115 Dublin Woods Drive	27513	Dormitory	
Jim	Phelps	1406 Oakcreek Drive	14450	Two-family house	

Based on the suppression data, you would not want to send any mailings to Bill Jones, as he already owns a home. If you were selling home-improvement services, you probably wouldn't want to send mailings to dormitory dweller Mary Wise.

Now let's look at another data append for the consumer records above. Using a third-party demographic database, we'll enhance our data for each consumer:

First Name	Last Name	Street Address	ZIP	Age	Estimated Household Income	Dwelling Unit Size	Children
Bill	Jones	1324 Milton Road	27712	57	\$50,000	3,000	4
Mary	Wise	115 Dublin Woods Drive	27513	23	\$18,000	1,500	0
Jim	Phelps	1406 Oakcreek Drive	14450	54	\$75,000	4,750	2

With this data, you can now target these people better. For example, based on income and dwelling size, you might offer home-improvement services to Bill and Jim, but storage rentals to Mary.

DATA MANAGEMENT IN ACTION: BELL MOBILITY

One DataFlux customer, Bell Mobility, is using DataFlux technology to manage data within its database marketing efforts. With over 4 million wireless subscribers, Bell Mobility supports a complete range of innovative wireless communications solutions: PCS and cellular, Web browsing and data, two-way messaging, paging, and airline passenger communications services.

To gain new subscribers, Bell Mobility uses telemarketing and direct mail campaigns to contact and engage prospective customers. The company typically receives prospect lists from external sources—such as Dun & Bradstreet or infoCANADA—requiring the company to check for duplicate entries across new and existing databases.

Traditionally, the company used SQL code to create the matching procedures that helped eliminate duplicate entries. However, this was a time-consuming and labor-intensive process, and Bell Mobility needed a better tool to quickly build an effective data quality process for contact lists.

THE SOLUTION: DATAFLUX TECHNOLOGY

Bell Mobility looked for a data management solution that could help it enhance response rates from its telemarketing and direct mail efforts through better contact information. With a highly accurate list of prospects, Bell Mobility could reach the right customer more often—and build more efficient database marketing campaigns.

The company chose dfPower Studio from DataFlux, an end-user solution that gives both business analysts and database administrators the ability to create, manage, and enforce data quality initiatives across the enterprise. With dfPower Studio, Bell Mobility can create custom data verification routines to help them identify and eliminate duplicate prospect records—and enhance the quality of mailing and telemarketing lists.

THE RESULTS

By performing extensive data quality routines on new and existing contact information, Bell Mobility provided more reliable data to its telemarketing and direct mail groups, allowing them to contact more people than in previous campaigns. With better prospect data, the company was able to add over \$1 million CAD (approx. \$730,000 USD) in annualized revenue through more effective database marketing campaigns. Additionally, dfPower Studio provided the automation that Bell Mobility needed to save time and labor costs.

The effectiveness of our database marketing programs at companies like Bell Mobility is based to a large degree on the quality of data available. With dfPower Studio, the company can create cleaner and more accurate lists that drive marketing programs. If the marketing team can provide a telemarketer with a better group of names, then they spend more time making calls to valid potential clients.

CONCLUSION

In a perfect world, every piece of communication that a marketer develops would reach the intended recipient. In the format that they prefer. And in a timeframe that is useful for them. Unfortunately, marketing in the real world is often complicated by the inability to develop and maintain consistent, accurate and reliable data sources that serve as the foundation for database marketing efforts.

As a result, many marketing databases cannot adequately support their intended activities. However, you can improve the value of your data. By following a proven, carefully ordered process of data management that follows the four building blocks (profiling, quality, integration and augmentation), you can achieve valuable results with limited time and resources.

And you can periodically check the condition of your data over time to ensure that your data does not degrade as more information hits the system. Or run data cleansing and verification processes in real-time, helping ensure that new data entering your sources is usable and actionable. With these processes in place, you can develop the data you need to conduct effective marketing campaigns that deliver impressive results.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

John Walker
Bell Mobility
5099 Creebank Rd
Mississauga, ON L4W 5N2
Work Phone: (905) 282-4220
Fax: (905)282-3525
Email: jowalker@mobility.com
Web: www.mobility.com

Gavin Day
DataFlux, a SAS company
3001 Weston Parkway, Suite 300
Cary, NC 27513
Work Phone: (919) 674-2153
Fax: (919) 678-8330
Email: gavin.day@dataflux.com
Web: www.dataflux.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

DataFlux and all other DataFlux Corporation product or service names are registered trademarks or trademarks of, or licensed to, DataFlux Corporation in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.